Solutions of the matrix inequalities $BXB^* \leqslant^- A$ in the minus partial ordering and $BXB^* \leqslant^{\mathrm{L}} A$ in the Löwner partial ordering

Yongge Tian

CEMA, Central University of Finance and Economics, Beijing 100081, China

Abstract. Two matrices A and B of the same size are said to satisfy the minus partial ordering, denoted by $B \leqslant^- A$, iff the rank subtractivity equality $\mathrm{rank}(\,A - B\,) = \mathrm{rank}(A) - \mathrm{rank}(B)$ holds; two complex Hermitian matrices A and B of the same size are said to satisfy the Löwner partial ordering, denoted by $B \leqslant^{\mathrm{L}} A$, iff the difference $A - B$ is nonnegative definite. In this note, we establish general solution of the inequality $BXB^* \leqslant^- A$ induced from the minus partial ordering, and general solution of the inequality $BXB^* \leqslant^{\mathrm{L}} A$ induced from the Löwner partial ordering, respectively, where $(\cdot)^*$ denotes the conjugate transpose of a complex matrix. As consequences, we give closed-form expressions for the shorted matrices of A relative to the range of B in the minus and Löwner partial orderings, respectively, and show that these two types of shorted matrices in fact are the same.

Mathematics Subject Classifications: 15A03; 15A09; 15A24; 15B57

Keywords: Minus partial ordering; Löwner partial ordering; Hermitian matrix; matrix function; matrix equation; Moore–Penrose inverse; shorted matrix; rank; inertia

1 Introduction

Throughout this note, let $\mathbb{C}^{m\times n}$ and $\mathbb{C}^{m}_{\mathrm{H}}$ denote the collections of all $m \times n$ complex matrices and all $m\times m$ complex Hermitian matrices, respectively; the symbols A^*, $r(A)$ and $\mathscr{R}(A)$ stand for the conjugate transpose, the rank and the range (column space) of a matrix $A \in \mathbb{C}^{m\times n}$, respectively; I_m denotes the identity matrix of order m; $[\,A,\,B\,]$ denotes a row block matrix consisting of A and B. The Moore–Penrose inverse of a matrix $A \in \mathbb{C}^{m\times n}$, denoted by $A^{\dagger}$, is defined to be the unique matrix $X \in \mathbb{C}^{n\times m}$ satisfying the matrix equations

$$\text{(i) } AXA = A, \quad \text{(ii) } XAX = X, \quad \text{(iii) } (AX)^* = AX, \quad \text{(iv) } (XA)^* = XA.$$

Further, let $E_A = I_m - AA^{\dagger}$ and $F_A = I_n - A^{\dagger}A$, both of which are orthogonal projectors and their ranks are given by $r(E_A) = m - r(A)$ and $r(F_A) = n - r(A)$. A well-known property of the Moore–Penrose inverse is $(A^{\dagger})^* = (A^*)^{\dagger}$. Hence, if $A = A^*$, then both $A^{\dagger} = (A^{\dagger})^*$ and $AA^{\dagger} = A^{\dagger}A$ hold. The inertia of a matrix $A \in \mathbb{C}^{m}_{\mathrm{H}}$ is defined to be the triplet $\mathrm{In}(A) = \{\, i_+(A),\, i_-(A),\, i_0(A)\,\}$, where $i_+(A)$, $i_-(A)$ and $i_0(A)$ are the numbers of the positive, negative and zero eigenvalues of A counted with multiplicities, respectively. For a matrix $A \in \mathbb{C}^{m}_{\mathrm{H}}$, both $r(A) = i_+(A) + i_-(A)$ and $i_0(A) = m - r(A)$ hold.

The definitions of two well-known partial orderings on matrices of the same size are given below.

Definition 1.1 (a) Two matrices $A,\, B \in \mathbb{C}^{m\times n}$ are said to satisfy the minus partial ordering, denoted by $B \leqslant^- A$, iff the rank subtractivity equality $r(\,A - B\,) = r(A) - r(B)$ holds, or equivalently, both $\mathscr{R}(A - B) \cap \mathscr{R}(B) = \{0\}$ and $\mathscr{R}(A^* - B^*) \cap \mathscr{R}(B^*) = \{0\}$ hold.

(b) Two matrices $A,\, B \in \mathbb{C}^{m}_{\mathrm{H}}$ are said to satisfy the Löwner partial ordering, denoted by $B \leqslant^{\mathrm{L}} A$, iff the difference $A - B$ is nonnegative definite, or equivalently, $A - B = UU^*$ for some matrix U.

In this note, we consider the following two matrix inequalities

$$BXB^* \leqslant^- A, \tag{1.1}$$

$$BXB^* \leqslant^{\mathrm{L}} A \tag{1.2}$$

induced from the minus and Löwner partial orderings, and examine the relations of their solutions, where $A \in \mathbb{C}^{m}_{\mathrm{H}}$ and $B \in \mathbb{C}^{m\times n}$ are given, and $X \in \mathbb{C}^{n}_{\mathrm{H}}$ is unknown. This consideration is motivated by some recent work on rank and inertia optimizations of $A - BXB^*$ in [7, 13, 14]. We shall derive general solutions of (1.1) and (1.2) by using the given matrices and their generalized inverses, and then discuss some algebraic properties of these solutions. In particular, we give solutions of the following constrained rank and Löwner partial ordering optimization problems

$$\max_{BXB^* \leqslant^- A} r(BXB^*), \qquad \min_{BXB^* \leqslant^- A} r(\,A - BXB^*\,), \tag{1.3}$$

E-mail Address: yongge.tian@gmail.com

$$\max_{\leqslant^{\mathrm{L}}}\{\, BXB^* \mid BXB^* \leqslant^{\mathrm{L}} A \,\}, \qquad \min_{\leqslant^{\mathrm{L}}}\{\, A - BXB^* \mid BXB^* \leqslant^{\mathrm{L}} A \,\}. \tag{1.4}$$

Eqs. (1.1) and (1.2) are equivalent to determining elements in the following matrix sets:

$$\mathcal{S}_1 = \{\, Z \in \mathbb{C}_{\mathrm{H}}^{m} \mid Z \leqslant^{-} A, \ \ \mathscr{R}(Z) \subseteq \mathscr{R}(B) \,\}, \tag{1.5}$$

$$\mathcal{S}_2 = \{\, Z \in \mathbb{C}_{\mathrm{H}}^{m} \mid Z \leqslant^{\mathrm{L}} A, \ \ \mathscr{R}(Z) \subseteq \mathscr{R}(B) \,\}. \tag{1.6}$$

The matrices Z in (1.5) and (1.6) can be regarded as two constrained approximations of the matrix A in partial orderings. In particular, a matrix $Z \in \mathcal{S}_1$ that has the maximal possible rank is called a shorted matrix of A relative to $\mathscr{R}(B)$ in the minus partial ordering (see [9, 11]); while the maximal matrix in $\mathcal{S}_2$ is called a shorted matrix of A relative to $\mathscr{R}(B)$ in the Löwner partial ordering (see [1, 2]). Our approaches to (1.1)–(1.4) link some previous and recent work in [1, 2, 3, 4, 5, 9, 10, 11] on shorted matrices of A relative to given subspaces in partial orderings, and some recent work on the rank and inertia of the matrix function $A - BXB^*$ in [7, 13, 14]. It is obvious that there always exists a matrix X that satisfies (1.1), say, $X = 0$. Hence, what we need to do is to derive a general expression of X that satisfies (1.1). Eq. (1.2) may have no solutions unless the given matrices A and B in (1.2) satisfy certain conditions.

This note is organized as follows. In Section 2, we present some known results on ranks and inertias of matrices and matrix equations, and then solve two homogeneous matrix equations with symmetric patterns. In Section 3, we use the results obtained in Section 2 to derive the general solution of (1.1), and give an analytical expression for the shorted matrix of A relative to $\mathscr{R}(B)$ in the minus partial ordering. In Section 4, we derive necessary and sufficient conditions for (1.2) to have a solution, and then give the general solution of (1.2). We show in Section 5 an interesting fact that the shorted matrices of A relative to $\mathscr{R}(B)$ in the minus and Löwner partial orderings are the same.

2 Preliminary results

In order to characterize matrix equalities that involve the Moore–Penrose inverses, we need the following rank and inertia expansion formulas.

Lemma 2.1 ([8]) *Let $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{m\times k}$ and $C \in \mathbb{C}^{l\times n}$ be given. Then, the following rank expansion formulas hold*

$$r[\, A, \, B \,] = r(A) + r(E_A B) = r(B) + r(E_B A), \tag{2.1}$$

$$r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r(CF_A) = r(C) + r(AF_C), \tag{2.2}$$

$$r\begin{bmatrix} A & B \\ C & CA^{\dagger}B \end{bmatrix} = r\begin{bmatrix} A \\ C \end{bmatrix} + r[\, A, \, B \,] - r(A). \tag{2.3}$$

Lemma 2.2 ([13]) *Let $A \in \mathbb{C}_{\mathrm{H}}^{m}$, $B \in \mathbb{C}^{m\times n}$, and $D \in \mathbb{C}_{\mathrm{H}}^{n}$. Then, the following inertia expansion formulas hold*

$$i_{\pm}\begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix} = r(B) + i_{\pm}(E_B A E_B), \tag{2.4}$$

$$i_{\pm}\begin{bmatrix} A & B \\ B^* & D \end{bmatrix} = i_{\pm}(A) + i_{\pm}(\, D - B^*A^{\dagger}B \,) \quad \textit{for } \mathscr{R}(B) \subseteq \mathscr{R}(A). \tag{2.5}$$

In order to solve (1.1) and (1.2), we also need the following results on solvability conditions and general solutions of two simple linear matrix equations.

Lemma 2.3 *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{m\times p}$ be given. Then, the following hold.*

(a) [12] *The matrix equation $AX = B$ is consistent if and only if $\mathscr{R}(B) \subseteq \mathscr{R}(A)$. In this case, the general solution can be written as $X = A^{\dagger}B + F_A U$, where $U \in \mathbb{C}^{n\times p}$ is arbitrary.*

(b) [6] *Under $B \in \mathbb{C}^{m\times n}$, the matrix equation $AX = B$ has a solution $0 \leqslant^{\mathrm{L}} X \in \mathbb{C}_{\mathrm{H}}^{n}$ if and only if $\mathscr{R}(B) \subseteq \mathscr{R}(A)$, $AB^* \geqslant^{\mathrm{L}} 0$ and $r(AB^*) = r(B)$. In this case, the general nonnegative definite solution can be written as*

$$X = B^*(AB^*)^{\dagger}B + F_A U F_A, \tag{2.6}$$

where $0 \leqslant^{\mathrm{L}} U \in \mathbb{C}_{\mathrm{H}}^{n}$ is arbitrary.

Lemma 2.4 *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^m_{\mathrm{H}}$ be given. Then, the following hold.*

(a) [4] *The matrix equation*

$$AXA^* = B \tag{2.7}$$

has a solution $X \in \mathbb{C}^n_{\mathrm{H}}$ if and only if $\mathscr{R}(B) \subseteq \mathscr{R}(A)$, or equivalently, $AA^\dagger B = B$.

(b) [13] *Under $X \in \mathbb{C}^n_{\mathrm{H}}$, the general Hermitian solution of (2.7) can be written in the following two forms*

$$X = A^\dagger B(A^\dagger)^* + U - A^\dagger AUA^\dagger A, \tag{2.8}$$
$$X = A^\dagger B(A^\dagger)^* + F_A V + V^* F_A, \tag{2.9}$$

respectively, where $U \in \mathbb{C}^n_{\mathrm{H}}$ and $V \in \mathbb{C}^{n\times n}$ are arbitrary.

Lemma 2.5 *Let $P \in \mathbb{C}^{m\times n}$ and $Q \in \mathbb{C}^{m\times k}$ be given. Then, the general solutions $X \in \mathbb{C}^n_{\mathrm{H}}$ and $Y \in \mathbb{C}^k_{\mathrm{H}}$ of the matrix equation*

$$PXP^* = QYQ^* \tag{2.10}$$

can be written as

$$X = X_1WX_1^* + X_2, \qquad Y = Y_1WY_1^* + Y_2, \tag{2.11}$$

where $W \in \mathbb{C}^m_{\mathrm{H}}$ is arbitrary, and $X_1 \in \mathbb{C}^{n\times m}$, $X_2 \in \mathbb{C}^n_{\mathrm{H}}$, $Y_1 \in \mathbb{C}^{k\times m}$ and $Y_2 \in \mathbb{C}^k_{\mathrm{H}}$ are the general solutions of the following matrix equations

$$PX_1 = QY_1, \quad PX_2P^* = 0, \quad QY_2Q^* = 0, \tag{2.12}$$

or alternatively, the general solution of (2.10) can be written in the following pair of parametric form

$$X = \widehat{I}_n F_H U F_H \widehat{I}_n^* + U_1 - P^\dagger PU_1P^\dagger P, \tag{2.13}$$
$$Y = \widetilde{I}_k F_H U F_H \widetilde{I}_k^* + U_2 - Q^\dagger QU_2Q^\dagger Q, \tag{2.14}$$

where $H = [\,P,\, -Q\,]$, $\widehat{I}_n = [\,I_n,\, 0\,]$, $\widetilde{I}_k = [\,0,\, I_k\,]$, and $U \in \mathbb{C}^{n+k}_{\mathrm{H}}$, $U_1 \in \mathbb{C}^n_{\mathrm{H}}$ and $U_2 \in \mathbb{C}^k_{\mathrm{H}}$ are arbitrary.

Proof It is easy to verify that the pair of matrices X and Y in (2.11) are both Hermitian. Substituting the pair of matrices into (2.10) gives

$$PXP^* = PX_1WX_1^*P^* = QY_1WY_1^*Q^* = QYQ^*,$$

which shows that (2.11) satisfies (2.10). Also, assume that X_0 and Y_0 are any pair of solutions of (2.10), and set

$$W = (PX_0P^*)^\dagger = (QY_0Q^*)^\dagger, \quad X_1 = P^\dagger PX_0P^*, \quad Y_1 = Q^\dagger QY_0Q^*,$$
$$X_2 = X_0 - P^\dagger PX_0P^\dagger P, \qquad Y_2 = Y_0 - Q^\dagger QY_0Q^\dagger Q.$$

Then, (2.11) reduces to

$$\begin{aligned} X &= P^\dagger PX_0P^*(PX_0P^*)^\dagger(P^\dagger PX_0P^*)^* + X_0 - P^\dagger PX_0P^\dagger P \\ &= P^\dagger(PX_0P^*)(PX_0P^*)^\dagger(PX_0P^*)(P^\dagger)^* + X_0 - P^\dagger PX_0P^\dagger P \\ &= P^\dagger PX_0P^\dagger P + X_0 - P^\dagger PX_0P^\dagger P = X_0, \\ Y &= Q^\dagger QY_0Q^*(QY_0Q^*)^\dagger(Q^\dagger QY_0Q^*)^* + Y_0 - Q^\dagger QY_0Q^\dagger Q \\ &= Q^\dagger(QY_0Q^*)(QY_0Q^*)^\dagger(QY_0Q^*)(Q^\dagger)^* + Y_0 - Q^\dagger QY_0Q^\dagger Q \\ &= Q^\dagger QY_0Q^\dagger Q + Y_0 - Q^\dagger QY_0Q^\dagger Q = Y_0, \end{aligned}$$

that is, any pair of solutions of (2.10) can be represented by (2.11). Thus, (2.11) is the general solution of (2.10).

Solving the latter two equations in (2.12) by Lemma 2.4(b) yields the following general solutions

$$X_2 = U_1 - P^\dagger PU_1P^\dagger P, \qquad Y_2 = U_2 - Q^\dagger QU_2Q^\dagger Q, \tag{2.15}$$

where $U_1 \in \mathbb{C}_{\mathrm{H}}^{n}$ and $U_2 \in \mathbb{C}_{\mathrm{H}}^{k}$ are arbitrary. To solve the first equation in (2.12), we rewrite it as $[\,P,\,-Q\,]\begin{bmatrix} X_1 \\ Y_1 \end{bmatrix} = 0$. Solving this equation by Lemma 2.3(a) gives the general solution $\begin{bmatrix} X_1 \\ Y_1 \end{bmatrix} = F_H V_1$, where V_1 is an arbitrary matrix. Hence, the general expressions of X_1 and Y_1 can be written as

$$X_1 = \widehat{I}_n F_H V_1, \quad Y_1 = \widetilde{I}_k F_H V_1. \tag{2.16}$$

Substituting (2.15) and (2.16) into (2.11) gives (2.13) and (2.14). $\square$

Lemma 2.6 *Let $B \in \mathbb{C}^{m\times n}$ and $A \in \mathbb{C}_{\mathrm{H}}^{m}$ be given. Then, the general solution $X \in \mathbb{C}_{\mathrm{H}}^{n}$ of the quadratic matrix equation*

$$(BXB^*)A(BXB^*) = BXB^* \tag{2.17}$$

can be expressed in the following parametric form

$$X = U(U^*B^*ABU)^{\dagger}U^* + V - B^{\dagger}BVB^{\dagger}B, \tag{2.18}$$

where $U \in \mathbb{C}^{n\times n}$ and $V \in \mathbb{C}_{\mathrm{H}}^{n}$ are arbitrary.

Proof Substituting (2.18) into BXB^* gives $BXB^* = BU(U^*B^*ABU)^{\dagger}U^*B^*$. It is easy to verify by the definition of the Moore–Penrose inverse that

$$\begin{aligned}(BXB^*)A(BXB^*) &= BU(U^*B^*ABU)^{\dagger}U^*B^*ABU(U^*B^*ABU)^{\dagger}U^*B^* \\ &= BU(U^*B^*ABU)^{\dagger}U^*B^* = BXB^*.\end{aligned}$$

Hence, (2.18) satisfies (2.17). On the other hand, for any Hermitian solution X_0 of (2.17), set $U = B^{\dagger}BX_0B^{\dagger}B$ and $V = X_0$ in (2.18). Then, (2.18) reduces to

$$\begin{aligned}X &= B^{\dagger}BX_0B^{\dagger}B(B^{\dagger}BX_0B^*ABX_0B^{\dagger}B)^{\dagger}B^{\dagger}BX_0B^{\dagger}B + X_0 - B^{\dagger}BX_0B^{\dagger}B \\ &= B^{\dagger}BX_0B^{\dagger}B(B^{\dagger}BX_0B^{\dagger}B)^{\dagger}B^{\dagger}BX_0B^{\dagger}B + X_0 - B^{\dagger}BX_0B^{\dagger}B \\ &= B^{\dagger}BX_0B^{\dagger}B + X_0 - B^{\dagger}BX_0B^{\dagger}B = X_0.\end{aligned}$$

This result indicates that all solutions of (2.17) can be represented through (2.18). Hence, (2.18) is the general solution of (2.17). $\square$

3 General solution of $BXB^* \leqslant^{-} A$

A well-known necessary and sufficient condition for the rank subtractivity equality in Definition 1.1 to hold is

$$r(\,A - B\,) = r(A) - r(B) \Leftrightarrow \mathscr{R}(B) \subseteq \mathscr{R}(A),\ \mathscr{R}(B^*) \subseteq \mathscr{R}(A^*) \text{ and } BA^{\dagger}B = B, \tag{3.1}$$

see [8]. Applying (3.1) to (1.1), we can convert (1.1) to a system of matrix equations.

Lemma 3.1 *Eq.* (1.1) *is equivalent to the following system of matrix equations*

$$BXB^* = AYA, \quad (BXB^*)A^{\dagger}(BXB^*) = BXB^*, \tag{3.2}$$

where $Y \in \mathbb{C}_{\mathrm{H}}^{m}$ is an unknown matrix.

Proof From (3.1), the minus partial order $BXB^* \leqslant^{-} A$ in (1.1) is equivalent to

$$\mathscr{R}(BXB^*) \subseteq \mathscr{R}(A) \ \text{ and } \ (BXB^*)A^{\dagger}(BXB^*) = BXB^*. \tag{3.3}$$

By Lemma 2.4(a), the first range inclusion in (3.3) holds if and only if the first matrix equation in (3.2) is solvable for Y. Thus, (3.2) and (3.3) are equivalent. $\square$

Theorem 3.2 *Let $A \in \mathbb{C}_{\mathrm{H}}^{m}$ and $B \in \mathbb{C}^{m\times n}$ be given, and $\mathcal{S}_1$ be as given in* (1.6). *Also define*

$$M = \begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix}, \quad H = [\,B,\,-A\,], \quad \widehat{I}_n = [\,I_n,\,0\,], \quad \widehat{B} = [\,B,\,0\,], \quad A_1 = E_BA, \quad B_1 = E_AB.$$

Then, the following hold.

(a) *The general Hermitian solution of the inequality*

$$BXB^* \leqslant^- A \tag{3.4}$$

can be written as

$$X = \widehat{I}_n F_H U(U^* F_H \widehat{B}^* A^\dagger \widehat{B} F_H U)^\dagger U^* F_H \widehat{I}_n^* + V - B^\dagger B V B^\dagger B, \tag{3.5}$$

where $U \in \mathbb{C}^{(m+n)\times(m+n)}$ *and* $V \in \mathbb{C}_{\mathrm{H}}^n$ *are arbitrary.*

(b) *The general expression of the matrices in* (1.5) *can be written as*

$$Z = \widehat{B} F_H U(U^* F_H \widehat{B}^* A^\dagger \widehat{B} F_H U)^\dagger U^* F_H \widehat{B}^*, \tag{3.6}$$

where $U \in \mathbb{C}^{(m+n)\times(m+n)}$ *is arbitrary. The global maximal and minimal inertias and ranks of* Z *in* (3.6) *and the corresponding* $A - Z$ *are given by*

$$\max_{Z\in\mathcal{S}_1} i_{\pm}(Z) = i_{\mp}(M) + i_{\pm}(A) - r[\,A,\, B\,], \tag{3.7}$$

$$\max_{Z\in\mathcal{S}_1} r(Z) = r(M) + r(A) - 2r[\,A,\, B\,], \tag{3.8}$$

$$\min_{Z\in\mathcal{S}_1} i_{\pm}(\,A - Z\,) = r[\,A,\, B\,] - i_{\mp}(M), \tag{3.9}$$

$$\min_{Z\in\mathcal{S}_1} r(\,A - Z\,) = 2r[\,A,\, B\,] - r(M). \tag{3.10}$$

The shorted matrix of A *relative to* $\mathscr{R}(B)$*, denoted by* $\phi^-(\,A\,|\,B\,)$*, which is a matrix* Z *that satisfies* (3.8)*, is given by*

$$\phi^-(\,A\,|\,B\,) = \widehat{B} F_H (F_H \widehat{B}^* A^\dagger \widehat{B} F_H)^\dagger F_H \widehat{B}^*. \tag{3.11}$$

Proof Applying Lemma 2.5 to the first equation in (3.2), we obtain the general solutions of X and Y as follows

$$X = \widehat{I}_n F_H T F_H \widehat{I}_n^* + V - B^\dagger B V B^\dagger B, \quad Y = \widehat{I}_m F_H T F_H \widehat{I}_m^* + W - A^\dagger A W A^\dagger A, \tag{3.12}$$

where $T \in \mathbb{C}_{\mathrm{H}}^{m+n}$, $V \in \mathbb{C}_{\mathrm{H}}^n$ and $W \in \mathbb{C}_{\mathrm{H}}^m$ are arbitrary. Substituting (3.12) into the second equation in (3.2) leads to the following quadratic matrix equation

$$(\widehat{B} F_H T F_H \widehat{B}^*) A^\dagger (\widehat{B} F_H T F_H \widehat{B}^*) = \widehat{B} F_H T F_H \widehat{B}^*.$$

By Lemma 2.6, the general solution of this quadratic matrix equation is given by

$$T = U(U^* F_H \widehat{B}^* A^\dagger \widehat{B} F_H U)^\dagger U^* + W_1 - (\widehat{B} F_H)^\dagger (\widehat{B} F_H) W_1 (\widehat{B} F_H)^\dagger (\widehat{B} F_H),$$

where $U \in \mathbb{C}^{(m+n)\times(m+n)}$ and $W_1 \in \mathbb{C}_{\mathrm{H}}^{m+n}$ are arbitrary. Substituting this T into the matrix X in (3.12) gives

$$\begin{aligned} X &= \widehat{I}_n F_H U(U^* F_H \widehat{B}^* A^\dagger \widehat{B} F_H U)^\dagger U^* F_H \widehat{I}_n^* \\ &\quad + [\,\widehat{I}_n F_H W_1 F_H \widehat{I}_n^* - \widehat{I}_n F_H (\widehat{B} F_H)^\dagger (\widehat{B} F_H) W_1 (\widehat{B} F_H)^\dagger (\widehat{B} F_H) F_H \widehat{I}_n^*\,] + V - B^\dagger B V B^\dagger B. \end{aligned} \tag{3.13}$$

It is easy to verify from $B\widehat{I}_n = \widehat{B}$ that

$$\begin{aligned} &B[\,\widehat{I}_n F_H W_1 F_H \widehat{I}_n^* - \widehat{I}_n F_H (\widehat{B} F_H)^\dagger (\,\widehat{B} F_H) W_1 (\widehat{B} F_H\,)^\dagger (\widehat{B} F_H) F_H \widehat{I}_n^*\,] B^* \\ &= \widehat{B} F_H W_1 F_H \widehat{B}^* - (\widehat{B} F_H)(\widehat{B} F_H)^\dagger (\,\widehat{B} F_H) W_1 (\widehat{B} F_H\,)^\dagger (\widehat{B} F_H)(\widehat{B} F_H)^* = 0. \end{aligned}$$

This fact shows that the second term on the right-hand side of (3.13) is a solution to $BXB^* = 0$. Also, note from Lemma 2.4(b) that $V - B^\dagger B V B^\dagger B$ is the general solution to $BXB^* = 0$. Hence, the second term on the right-hand side of (3.13) can be represented by the third term of the same side, so that (3.13) reduces to (3.5).

Substituting (3.5) into BXB^* gives

$$Z = BXB^* = \widehat{B} F_H U(\,U^* F_H \widehat{B}^* A^\dagger \widehat{B} F_H U\,)^\dagger U^* F_H \widehat{B}^*, \tag{3.14}$$

as required for (3.6). Note further that this Z satisfies

$$(U^*F_H\widehat{B}^*A^\dagger)Z(A^\dagger\widehat{B}F_HU) = U^*F_H\widehat{B}^*A^\dagger\widehat{B}F_HU. \tag{3.15}$$

Both (3.14) and (3.15) imply

$$i_{\pm}(Z) = i_{\pm}(U^*F_H\widehat{B}^*A^\dagger\widehat{B}F_HU) \leqslant i_{\pm}(\, F_H\widehat{B}^*A^\dagger\widehat{B}F_H\,) \tag{3.16}$$

and

$$\max_{Z\in\mathcal{S}_1} i_{\pm}(Z) = i_{\pm}(F_H\widehat{B}^*A^\dagger\widehat{B}F_H), \quad \max_{Z\in\mathcal{S}_1} r(Z) = r(F_H\widehat{B}^*A^\dagger\widehat{B}F_H). \tag{3.17}$$

Recall that the inertia of a Hermitian matrix does not change under Hermitian congruence operations. Applying (2.4) to $F_H\widehat{B}^*A^\dagger\widehat{B}F_H$ and simplifying by Hermitian congruence operations, we obtain

$$\begin{aligned} i_{\pm}(F_H\widehat{B}^*A^\dagger\widehat{B}F_H) &= i_{\pm}\begin{bmatrix} \widehat{B}^*A^\dagger\widehat{B} & H^* \\ H & 0 \end{bmatrix} - r(H) \\ &= i_{\pm}\begin{bmatrix} B^*A^\dagger B & 0 & B^* \\ 0 & 0 & -A \\ B & -A & 0 \end{bmatrix} - r[\,A,\, B\,] \\ &= i_{\pm}\begin{bmatrix} 0 & \frac{1}{2}B^*A^\dagger A & B^* \\ \frac{1}{2}A^\dagger AB & 0 & -A \\ B & -A & 0 \end{bmatrix} - r[\,A,\, B\,] \\ &= i_{\pm}\begin{bmatrix} 0 & 0 & B^* \\ 0 & A & -A \\ B & -A & 0 \end{bmatrix} - r[\,A,\, B\,] \\ &= i_{\pm}\begin{bmatrix} 0 & 0 & B^* \\ 0 & A & 0 \\ B & 0 & -A \end{bmatrix} - r[\,A,\, B\,] \\ &= i_{\mp}\begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix} + i_{\pm}(A) - r[\,A,\, B\,]. \end{aligned} \tag{3.18}$$

Substituting (3.18) into (3.17) leads to (3.7) and (3.8). Also, note that

$$\min_{X\in\mathcal{S}_1} i_{\pm}(\,A - Z\,) = i_{\pm}(A) - \max_{X\in\mathcal{S}_1} i_{\pm}(Z).$$

Thus, (3.9) and (3.10) follow from (3.7) and (3.8). $\square$

4 General solution of $BXB^* \leqslant^{\rm L} A$

In this section, we derive an analytical expression for the general solution of (1.2) by using generalized inverses of matrices, and show some algebraic properties of the solution.

Theorem 4.1 *Let $A \in \mathbb{C}_{\rm H}^{m}$ and $B \in \mathbb{C}^{m\times n}$ be given, and let $\mathcal{S}_2$ be as given in* (1.6)*. Then, the following hold.*

(a) *There exists an $X \in \mathbb{C}_{\rm H}^{n}$ such that*

$$BXB^* \leqslant^{\rm L} A \tag{4.1}$$

if and only if

$$E_BAE_B \geqslant^{\rm L} 0 \quad \textit{and} \quad r(E_BAE_B) = r(E_BA), \tag{4.2}$$

or equivalently,

$$i_{+}\begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix} = r[\,A,\, B\,] \quad \textit{and} \quad i_{-}\begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix} = r(B). \tag{4.3}$$

In this case, the general Hermitian solution of (4.1) *can be written in the following parametric form*

$$X = B^\dagger A(B^\dagger)^* - B^\dagger AE_B(E_BAE_B)^\dagger E_BA(B^\dagger)^* - UU^* + F_BV + V^*F_B, \tag{4.4}$$

where $U,\ V \in \mathbb{C}^{n\times n}$ are arbitrary. Correspondingly, the general expression of the matrices in $\mathcal{S}_2$ can be written as

$$Z = A - AE_B(E_BAE_B)^\dagger E_BA - BUU^*B^*. \tag{4.5}$$

(b) *Under (4.2), the shorted matrix of A relative to $\mathscr{R}(B)$, denoted by $\phi^{\rm L}(\,A\,|\,B\,)$, which is the maximizer in $\mathcal{S}_2$, can uniquely be written as*

$$\phi^{\rm L}(\,A\,|\,B\,) = A - AE_B(E_BAE_B)^{\dagger}E_BA. \tag{4.6}$$

The rank and inertia of $\phi^{\rm L}(\,A\,|\,B\,)$ and $A - \phi^{\rm L}(\,A\,|\,B\,)$ satisfy

$$i_{+}[\,\phi^{\rm L}(\,A\,|\,B\,)\,] = i_{+}(A) + r(B) - r[\,A,\,B\,], \tag{4.7}$$

$$i_{-}[\,\phi^{\rm L}(\,A\,|\,B\,)\,] = i_{-}(A), \tag{4.8}$$

$$i_{+}[\,A - \phi^{\rm L}(\,A\,|\,B\,)\,] = r[\,A - \phi^{\rm L}(\,A\,|\,B\,)\,] = r[\,A,\,B\,] - r(B). \tag{4.9}$$

Proof It is obvious that (4.1) is equivalent to

$$BXB^{*} = A - YY^{*} \tag{4.10}$$

for some matrix Y. In other words, (4.1) can be relaxed to a matrix equation with two unknown matrices. From Lemma 2.4(a), (4.10) is solvable for $X \in \mathbb{C}_{\rm H}^{n}$ if and only if $E_B(\,A - YY^{*}\,) = 0$, that is,

$$E_BYY^{*} = E_BA. \tag{4.11}$$

From Lemma 2.3(b), (4.11) is solvable for YY^{*} if and only if $E_BAE_B \succcurlyeq^{\rm L} 0$ and $r(E_BAE_B) = r(E_BA)$, establishing (4.2), which is further equivalent to (4.3) by (2.1) and (2.4). In this case, the general nonnegative definite solution of (4.11) can be written as

$$YY^{*} = AE_B(E_BAE_B)^{\dagger}E_BA + BB^{\dagger}WBB^{\dagger}, \tag{4.12}$$

where $0 \preccurlyeq^{\rm L} W \in \mathbb{C}_{\rm H}^{m}$ is arbitrary. Substituting the YY^{*} into (4.10) gives

$$BXB^{*} = A - AE_B(E_BAE_B)^{\dagger}E_BA - BB^{\dagger}WBB^{\dagger}. \tag{4.13}$$

By Lemma 2.4(b), the general Hermitian solution of (4.13) can be written as

$$X = B^{\dagger}A(B^{\dagger})^{*} - B^{\dagger}AE_B(E_BAE_B)^{\dagger}E_BA(B^{\dagger})^{*} - B^{\dagger}W(B^{\dagger})^{*} + F_BV + V^{*}F_B, \tag{4.14}$$

where $V \in \mathbb{C}^{n\times n}$ is arbitrary. Replacing the matrix $0 \preccurlyeq^{\rm L} B^{\dagger}W(B^{\dagger})^{*} \in \mathbb{C}_{\rm H}^{n}$ in (4.14) with a general matrix $0 \preccurlyeq^{\rm L} U \in \mathbb{C}_{\rm H}^{n}$ yields (4.4), which is also the general Hermitian solution of (4.1). Substituting (4.4) into BXB^{*} gives (4.5).

Eq. (4.6) follows from (4.5) by noticing $BUU^{*}B^{*} \succcurlyeq^{\rm L} 0$.

It follows from (4.2) that $\mathscr{R}(E_BAE_B) = \mathscr{R}(E_BA)$. In this case, applying (2.5) to (4.6) and simplifying by Hermitian congruence transformations, we obtain

$$\begin{aligned}
i_{\pm}[\,\phi^{\rm L}(\,A\,|\,B\,)\,] &= i_{\pm}[\,A - AE_B(E_BAE_B)^{\dagger}E_BA\,]\\
&= i_{\pm}\begin{bmatrix} E_BAE_B & E_BA \\ AE_B & A \end{bmatrix} - i_{\pm}(E_BAE_B)\\
&= i_{\pm}\begin{bmatrix} 0 & 0 \\ 0 & A \end{bmatrix} - i_{\pm}(E_BAE_B)\\
&= i_{\pm}(A) - i_{\pm}(E_BAE_B),\\
i_{\pm}[\,A - \phi^{\rm L}(\,A\,|\,B\,)\,] &= i_{\pm}[\,AE_B(E_BAE_B)^{\dagger}E_BA\,]\\
&= i_{\pm}\begin{bmatrix} -E_BAE_B & E_BA \\ AE_B & 0 \end{bmatrix} - i_{\mp}(E_BAE_B)\\
&= i_{\pm}\begin{bmatrix} 0 & E_BA \\ AE_B & 0 \end{bmatrix} - i_{\mp}(E_BAE_B)\\
&= r(E_BA) - i_{\mp}(E_BAE_B).
\end{aligned}$$

Hence, we further find from that (2.1) and (4.2) that

$$\begin{aligned}
i_{+}[\phi^{\rm L}(\,A\,|\,B\,)] &= i_{+}(A) - i_{+}(E_BAE_B) = i_{+}(A) - r(E_BA) = i_{+}(A) + r(B) - r[\,A,\,B\,],\\
i_{-}[\,\phi^{\rm L}(\,A\,|\,B\,)\,] &= i_{-}(A) - i_{-}(E_BAE_B) = i_{-}(A),\\
i_{+}[\,A - \phi^{\rm L}(\,A\,|\,B\,)\,] &= r(E_BA) - i_{-}(E_BAE_B) = r(E_BA) = r[\,A,\,B\,] - r(B),\\
i_{-}[\,A - \phi^{\rm L}(\,A\,|\,B\,)\,] &= r(E_BA) - i_{+}(E_BAE_B) = 0,
\end{aligned}$$

establishing (4.7)–(4.9). $\square$

5 An equality for the shorted matrices of A relative to $\mathscr{R}(B)$ in the minus and Löwner partial orderings

Since $\mathcal{S}_1$ and $\mathcal{S}_2$ in (1.5) and (1.6) are defined from different matrix inequalities, the two sets are not necessarily the same, as demonstrated in Theorems 3.2(b) and 4.1(a). However, they may have some common matrices. In this section, we show an interesting fact that the shorted matrices of A relative to $\mathscr{R}(B)$ in the minus and Löwner partial orderings are the same.

Theorem 5.1 *Let $A \in \mathbb{C}_{\rm H}^{m}$ and $B \in \mathbb{C}^{m\times n}$ be given, and $\mathcal{S}_1$ and $\mathcal{S}_2$ be as given in* (1.5) *and* (1.6)*. If* (4.1) *has a solution, then the two shorted matrices in $\mathcal{S}_1$ and $\mathcal{S}_2$ are the same, namely,*

$$\phi^{-}(\,A\,|\,B\,) = \phi^{\rm L}(\,A\,|\,B\,). \tag{5.1}$$

Proof Note from (3.11) and (4.6) that (5.1) holds if and only if

$$\widehat{B}F_H(F_H\widehat{B}^*A^{\dagger}\widehat{B}F_H)^{\dagger}F_H\widehat{B}^* = A - AE_B(E_BAE_B)^{\dagger}E_BA. \tag{5.2}$$

It is easy to derive from (2.2) that

$$r(\,\widehat{B}F_H\,) = r\begin{bmatrix} B \\ H \end{bmatrix} - r(H) = r(A) + r(B) - r[\,A,\; B\,]. \tag{5.3}$$

Under (4.2), (3.17) reduces to

$$r(\,F_H\widehat{B}^*A^{\dagger}\widehat{B}F_H\,) = r(M) + r(A) - 2r[\,A,\; B\,] = r(A) + r(B) - r[\,A,\; B\,]. \tag{5.4}$$

Both (5.3) and (5.4) imply that $\mathscr{R}(\,F_H\widehat{B}^*\,) = \mathscr{R}(\,F_H\widehat{B}^*A^{\dagger}\widehat{B}F_H\,)$. In this case, applying (2.5) to the difference of both sides of (5.2) and simplifying by elementary matrix operations, we obtain

$$\begin{aligned}
& r[\,A - AE_B(E_BAE_B)^{\dagger}E_BA - \widehat{B}F_H(F_H\widehat{B}^*A^{\dagger}\widehat{B}F_H)^{\dagger}F_H\widehat{B}^*\,] \\
&= r\begin{bmatrix} F_H\widehat{B}^*A^{\dagger}\widehat{B}F_H & F_H\widehat{B}^* \\ \widehat{B}F_H & A - AE_B(E_BAE_B)^{\dagger}E_BA \end{bmatrix} - r(F_H\widehat{B}^*A^{\dagger}\widehat{B}F_H) \\
&= r\begin{bmatrix} \widehat{B}^*A^{\dagger}\widehat{B} & \widehat{B}^* & H^* \\ \widehat{B} & A - AE_B(E_BAE_B)^{\dagger}E_BA & 0 \\ H & 0 & 0 \end{bmatrix} - 2r(H) - r(A) - r(B) + r[\,A,\; B\,] \quad \text{(by (2.4))} \\
&= r\begin{bmatrix} B^*A^{\dagger}B & 0 & B^* & B^* \\ 0 & 0 & 0 & -A \\ B & 0 & A - AE_B(E_BAE_B)^{\dagger}E_BA & 0 \\ B & -A & 0 & 0 \end{bmatrix} - r(A) - r(B) - r[\,A,\; B\,] \\
&= r\begin{bmatrix} B^*A^{\dagger}B & 0 & B^* & 0 \\ 0 & 0 & 0 & -A \\ B & 0 & A - AE_B(E_BAE_B)^{\dagger}E_BA & 0 \\ 0 & -A & 0 & 0 \end{bmatrix} - r(A) - r(B) - r[\,A,\; B\,] \\
&= r\begin{bmatrix} B^*A^{\dagger}B & B^* \\ B & A - AE_B(E_BAE_B)^{\dagger}E_BA \end{bmatrix} + r(A) - r(B) - r[\,A,\; B\,] \\
&= r\left(\begin{bmatrix} B^*A^{\dagger}B & B^* \\ B & A \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & AE_B(E_BAE_B)^{\dagger}E_BA \end{bmatrix}\right) + r(A) - r(B) - r[\,A,\; B\,] \\
&= r\begin{bmatrix} B^*A^{\dagger}B & B^* & 0 \\ B & A & AE_B \\ 0 & E_BA & E_BAE_B \end{bmatrix} - r(E_BAE_B) + r(A) - r(B) - r[\,A,\; B\,] \quad \text{(by (2.5))} \\
&= r\begin{bmatrix} B^*A^{\dagger}B & B^* & 0 \\ B & A & 0 \\ 0 & 0 & 0 \end{bmatrix} - r(E_BA) + r(A) - r(B) - r[\,A,\; B\,] \\
&= r\begin{bmatrix} B^*A^{\dagger}B & B^* \\ B & A \end{bmatrix} + r(A) - 2r[\,A,\; B\,] = 0 \quad \text{(by (2.3))},
\end{aligned}$$

which means that (5.2) is an equality. $\square$

The minus and Löwner partial orderings in Definition 1.1 can accordingly be defined for linear operators on a Hilbert space. Also, note that the results in this note are derived from some ordinary algebraic operations of the given matrices and their Moore–Penrose inverses. Hence, it is no doubt that most of the conclusions in this note can be extended to operator algebra, in which the Moore–Penrose inverses of linear operators were defined.

References

[1] W.N. Anderson, Jr., Shorted operators, SIAM J. Appl. Math. 20(1971), 522–525.

[2] W.N. Anderson, Jr. and G.E. Trapp, Shorted operators II, SIAM J. Appl. Math. 28(1975), 60–71.

[3] H. Goller, Shorted operators and rank decomposition matrices, Linear Algebra Appl. 81(1986), 207–236.

[4] J. Groß, A note on the rank-subtractivity ordering, Linear Algebra Appl. 289(1999), 151–160.

[5] R.E. Hartwig and G.P.H. Styan, On some characterizations of the "star" partial ordering for matrices and rank subtractivity, Linear Algebra Appl. 82(1986), 145–16l.

[6] C.G. Khatri and S.K. Mitra, Hermitian and nonnegative definite solutions of linear matrix equations, SIAM J. Appl. Math. 31(1976), 579–585.

[7] Y. Liu and Y. Tian, Hermitian-type of singular value decomposition for a pair of matrices and its applications, Numer. Linear Algebra Appl., DOI: 10.1002/nla.1825.

[8] G. Marsaglia and G.P.H. Styan, Equalities and inequalities for ranks of matrices, Linear and Multilinear Algebra 2(1974), 269–292.

[9] S.K. Mitra, The minus partial order and the shorted matrix, Linear Algebra Appl. 83(1986), 1–27.

[10] S.K. Mitra, P Bhimasankaram, S.B. Malik, Matrix Partial Orders, Shorted Operators and Applications, World Scientific Publishing, 2010.

[11] S.K. Mitra and M.L. Puri, Shorted matrices–an extended concept and some applications, Linear Algebra Appl. 42(1982), 57–79.

[12] R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc. 51(1955), 406–413.

[13] Y. Tian, Equalities and inequalities for inertias of Hermitian matrices with applications, Linear Algebra Appl. 433(2010), 263–296.

[14] Y. Tian and Y. Liu, Extremal ranks of some symmetric matrix expressions with applications, SIAM J. Matrix Anal. Appl. 28(2006), 890–905.